Renewable and Sustainable Energy Reviews 81 (2018) 912–928

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

Forecasting of photovoltaic power generation and model optimization: A review

Utpal Kumar Das[a], Kok Soon Tey[a,*], Mehdi Seyedmahmoudian[a], Saad Mekhilef[b], Moh Yamani Idna Idris[c], Willem Van Deventer[c], Bend Horan[c], Alex Stojcevski[d]

[a] Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia
[b] Power Electronics and Renewable Energy Research Laboratory (PEARL), Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia
[c] School of Engineering, Deakin University, Victoria 3216, Australia
[d] School of Software and Electrical Engineering, Swinburne University of Technology, Victoria 3122, Australia

ARTICLE INFO

Keywords:
PV power forecasting
Artificial intelligence
Machine-learning
Hybrid model
Optimization

ABSTRACT

To mitigate the impact of climate change and global warming, the use of renewable energies is increasing day by day significantly. A considerable amount of electricity is generated from renewable energy sources since the last decade. Among the potential renewable energies, photovoltaic (PV) has experienced enormous growth in electricity generation. A large number of PV systems have been installed in on-grid and off-grid systems in the last few years. The number of PV systems will increase rapidly in the future due to the policies of the government and international organizations, and the advantages of PV technology. However, the variability of PV power generation creates different negative impacts on the electric grid system, such as the stability, reliability, and planning of the operation, aside from the economic benefits. Therefore, accurate forecasting of PV power generation is significantly important to stabilize and secure grid operation and promote large-scale PV power integration. A good number of research has been conducted to forecast PV power generation in different perspectives. This paper made a comprehensive and systematic review of the direct forecasting of PV power generation. The importance of the correlation of the input-output data and the preprocessing of model input data are discussed. This review covers the performance analysis of several PV power forecasting models based on different classifications. The critical analysis of recent works, including statistical and machine-learning models based on historical data, is also presented. Moreover, the strengths and weaknesses of the different forecasting models, including hybrid models, and performance matrices in evaluating the forecasting model, are considered in this research. In addition, the potential benefits of model optimization are also discussed.

1. Introduction

Energy is a prime mover of economic growth and a key factor for the development of a country. Over the years, the demand for energy, especially electrical energy, has been increasing due to industrialization, modernization, and population growth, aside from people's increasing living standard. In the last century, electrical energy was generated mainly from fossil fuels (namely, oil, coal and natural gas) and currently still relies on them significantly. However, this issue has created environmental concerns about greenhouse gas (GHG) emissions, global warming, and climate change. The burning of fossil fuels releases carbon dioxide and some other toxic gases into the atmosphere, which contributes to global warming, as well as environmental pollution. Therefore, the modern world wants to increase the penetration of renewable energy aside from minimizing the use of fossil fuel in this sector.

During the United Nations Climate Change Conference (COP21) in 2015, also known as the Paris agreement, most of the countries agreed to make an effort to limit the global warming to ensure a livable world for the next generation. A joint presidential statement of the USA and

Abbreviations: PV, photovoltaic; NWP, numerical weather prediction; AI, artificial intelligence; AR, auto regressive; MA, moving average; ARMA, auto regressive moving average; ARIMA, AR integrated MA; ARMAX, ARMA exogenous; ANN, artificial neural network; SVM, support vector machine; SVR, support vector regression; HS, hybrid system; FS, fuzzy system; ANFIS, adaptive neuro fuzzy inference system; GA, genetic algorithm; GHG, greenhouse gas; IEA, international energy agency; MSE, mean square error; RMSE, root mean square error; nRMSE, normalized root mean square error; MAE, mean absolute error; MAPE, mean absolute percentage error; MRE, mean relative error; MBE, mean bias error

* Corresponding author.
E-mail address: koksoon@um.edu.my (K.S. Tey).

http://dx.doi.org/10.1016/j.rser.2017.08.017
Received 7 April 2017; Received in revised form 12 June 2017; Accepted 7 August 2017
Available online 31 August 2017
1364-0321/ © 2017 Elsevier Ltd. All rights reserved.

China on climate change highlights the new domestic policy commitments involving renewable energy and energy efficiency [1]. The European Union sets the production target of renewable energy to the domestic reduction of GHG emissions at least 40% by 2030 and 80% by 2050 (from a 1990 baseline). By 2030, 30% of the total energy will be produced from renewable energy sources and will achieve 100% by 2050 [2]. Similarly, most of the countries are committed to increase the use of renewable energy. Therefore, renewable energy has obtained the special attention of power production companies, energy policy makers, and governments.

Among the renewable energy resources, wind and solar energy are the more acceptable and promising sources due to their potential and availability. They have higher rate of success to penetration into the energy market. The sun, which is the source of solar energy, acts like a black body radiator with a surface temperature of 5800 K, which provides 1367 W/m^2 solar energy over the atmosphere [3–5]. A research study highlights that the earth receives approximately 1.8 × 10^{11} MW power from solar radiation at an instant [6]. However, the present world energy consumption requirement is less than the amount of energy received from the sun [7]. Therefore, solar energy has been obtaining much attention from energy players, investors, government, and international organizations in the last decade because of its several benefits including environmental and economic benefit.

Photovoltaic (PV) system has been achieving worldwide acceptance and is playing a significant role in providing clean and sustainable energy [8]. Therefore, PV power installation has experienced enormous growth over the last few years. Fig. 1 shows the installed worldwide capacity of solar PV power in the last decade. An exponential growth can be observed in the global PV power installed capacity from 2005 to 2015. The global capacity of PV power installation increased from 5.1 to 16.0 GW in 3 years (2005–2008) and continued to increase from 16.0 to 100.0 GW in the next 4 years (2008–2012). Overall, the global capacity of installed PV power has increased from 5.1 to 227.0 GW in the last 10 years. This growth will continue at a similar or higher rate in the future due to its technological benefit. Thus, the decrease of price (as low as $1.5/Wp for fixed-tilt systems) and the increase of the efficiency of the PV module will increase installations [9,10]. The high penetration of PV in electric systems has many economic benefits. The solar share has been reached annually in the electricity market, which is 7.8% in Italy and close to 6.5% in Germany and Greece; moreover, 22 countries have a percentage greater than 1% (IEA, 2016). According to IEA's highest projection, the installation of global solar PV power capacity could exceed 1700 GW by 2030 [11].

However, the generation of PV power fully depends on the uncertain and ungovernable meteorological factors, such as solar irradiance, atmospheric temperature, module temperature, wind pressure and direction, and humidity. The power output of a PV system dynamically changes with time due to the variability of environmental factors. Therefore, the accurate forecasting of PV power generation is considerably difficult. The unpredicted PV power output adversely affects the stability, reliability, and scheduling of the power system operation, aside from the economic benefit [12,13]. An accurate forecasting of the PV power generation can reduce the impact of PV power uncertainty on the grid, improve system reliability, maintain power quality, and increase the penetration level of the PV systems. Therefore, accurate forecasting of PV power generation is a great challenge for the researchers at this moment.

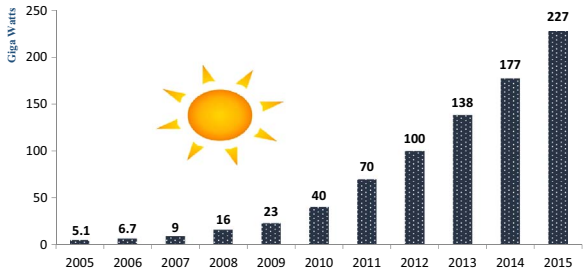


Fig. 1. Total installed PV power system from 2005 to 2015 worldwide [1].

A good number of research has been conducted to develop appropriate forecasting models in forecasting PV power generation with the targets of higher accuracy and minimum complexity with computational cost. These forecasting models are broadly classified into two categories: indirect and direct forecasting models. In the indirect forecasting model, the solar irradiance on different time scales have been forecasted by using various methods, including numerical weather prediction (NWP), image-based, statistical, and hybrid artificial neural network (ANN) based methods [14–21]. Therefore, the forecasted solar irradiance and others associated data are inputted in commercial PV simulation software, such as TRNSYSM, PVFORM, and HOMER [22], to forecast the PV power generation. However, in the direct forecasting model, PV power generation is forecasted directly using historical data samples, such as PV power output and associated meteorological data. Mitsuru *et al.* [23] have implemented direct and indirect methods to forecast the next-day power generation of a PV system, and showed that the direct method is better.

Several papers have reviewed the literature related to this field, focusing on different aspects. For example, in Ref. [24], the authors focused on wind speed/power forecasting and solar irradiance forecasting conducted by using ensemble methods. Gandomanin et al. [25] reviewed the literature on short-term forecasting of solar PV power output based on cloud cover influence. Antonanzas et al. [26] highlighted recent studies on solar power forecasting. However, PV power forecasting models were not extensively explored in their study. Wan et al. [27] analyzed different PV and solar forecasting techniques but did not review recent studies on PV power forecasting. On the other hand, Raza et al. [28] emphasized the PV power forecasting model beside the solar power forecasting model, which is insufficient. In addition, future studies should also consider the optimization of the PV power forecasting model.

Most of the recent studies in this field have focused on investigating direct PV power forecasting. Direct forecasting methods can achieve accurate forecasting of PV power generation. Therefore, a comprehensive literature review based on recent direct forecasting methods, including model development and optimization, should be conducted for new researchers in this field. This review will be helpful for the researchers in related fields, such as PV-integrated smart buildings, efficient energy management system, eclectic vehicle charging and smart grid.

The current work made a comprehensive and systematic review of the short-term direct forecasting models of PV power generation based on historical data. Section 2 of this paper presents a discussion on the importance of the correlation between input-output data and the preprocessing of input data to develop a PV power forecasting model. The classification of the forecasting models of PV power generation based on forecast horizon, historical data, and methodology is presented in Section 3. The strength and limitation of the different models, including statistical and artificial intelligence model based on historical data, are discussed in Section 4. The model optimization benefit is presented in Section 5. Section 6 introduces the performance matrices for evaluating the forecasting models. Section 7 presents an analysis of the recent works based on model performance and techniques. Finally, Section 8 summarizes and concludes the paper.

2. Input selection of the PV power forecasting model

Solar energy is one of most popular renewable energies, which comes from the sun in the form of solar irradiance. The solar cells made up of semiconductors in the PV module converts the solar irradiance to electricity through the photovoltaic effect. The PV power generation mainly depends on the amount of solar irradiance. In addition, other weather parameters, including atmospheric temperature, module tem-

perature, wind speed and direction, and humidity, are considered as potential parameters for estimating the PV power output.

A significant number of historical time series data of PV power output and corresponding meteorological variables are used to establish the forecasting model of PV power generation. The historical series data are divided in two groups: the training and testing data. The training data (almost 70% of the total data) are used for the learning of the model to forecast future values, whereas the testing data are used to validate the PV power forecasting model.

2.1. Correlation between PV power output and input variables

The variation of meteorological parameters depends on the geographical location, as well as the weather condition; thus, no similar impact of a meteorological parameter exists on the PV power generation at different geographical locations. Consequently, the correlation of the meteorological parameters and PV power output will not be the same in different locations. However, the performance of a forecasting model is highly dependent on the correlation of the input and output values of the model. In this case, the study of the correlation of the different meteorological inputs, such as solar irradiance, atmospheric temperature, module temperature, wind speed and direction, and humidity, with PV power output, is important. The correlation might be positive or negative. The strongly correlated input variables should be used as an input vector to the forecasting model, and the weakly correlated input vector data should be refused.

Fig. 2 shows the pattern of the solar irradiance and PV power output of a particular day. In a clear-sky day means a normal day, the PV power output is highly strongly matched with the solar irradiance curve. The PV power output is not highly strongly matched with the solar irradiance in an abnormal day, like cloudy or rainy day, but it is strongly matched. Therefore, a similar pattern is observed for PV power output and solar irradiance in any weather condition. Fig. 3 shows the strong positive correlation between solar irradiance and PV power output. Therefore, solar irradiance is an important input vector [29] in developing an appropriate PV power forecasting model due to its high correlation coefficient (R^2 = 0.988).

Fig. 4 highlights the pattern of the atmospheric temperature and PV power output in a particular day. The pattern of the atmospheric temperature and PV power almost matches the daylight period only. In the period of the absence of daylight, the PV power output is absent and no impact of atmospheric temperature exists on the PV power output. Fig. 5 indicates the correlation between the atmospheric temperature and the PV power output. The correlation is not strong, like irradiance, but not so weak as well. Other research studies also mentioned that, a strong correlation exists between atmospheric temperature and PV power output [30]. Therefore, the atmospheric temperature can be considered as a potential input vector in correctly establishing the forecasting model of the PV power generation.

The index data of module temperature, wind speed, humidity, and

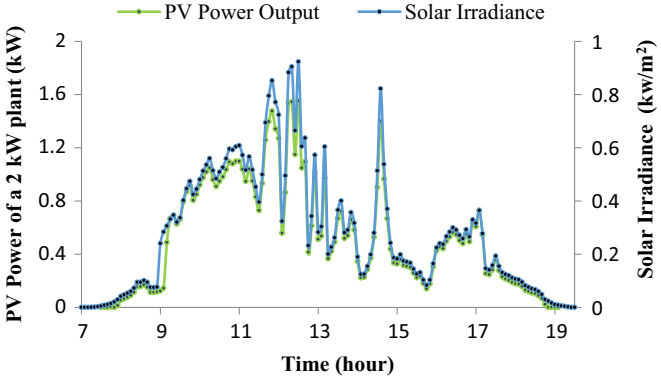


Fig. 2. Solar irradiance and PV power output pattern for a particular day.

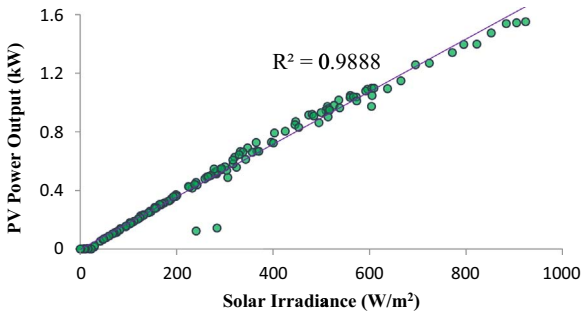


Fig. 3. Correlation between solar irradiance and PV power output.

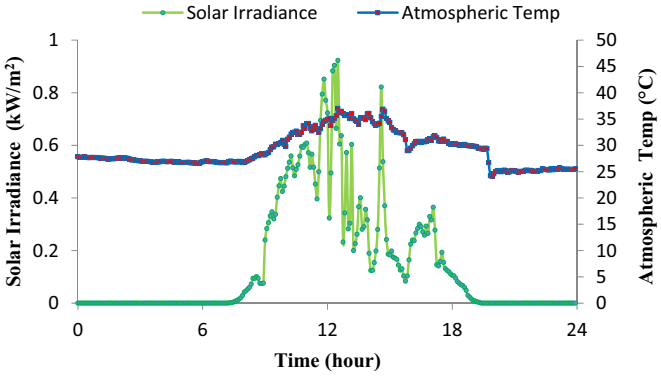


Fig. 4. Atmospheric temperature and PV power output pattern for a particular day.

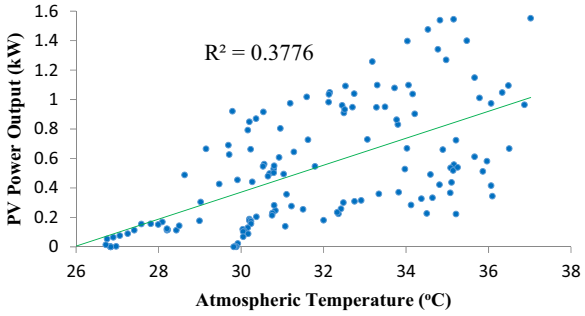


Fig. 5. Correlation between atmospheric temperature and PV power output.

cloud cover are also recorded in the data acquisition system of a PV plant. All the variables play a positive or negative role on the amount of PV power generation. The role of these variables depends on the geographical location; and the correlation of these variables with PV power output depends on the role of each variable. Suppose that the wind speed will not be able to contribute to change the amount of PV power generation in a geographical location with very low temperature. The accuracy of the forecasting model of the PV power generation may be enhanced using a large number of input vectors. However, the computational cost and complexity will also be increased due to the aggregation of the large number of input parameters. Therefore, designing a forecasting model with an optimal number of input vectors based on correlation is of utmost important.

From the above figures and discussion, a remarkably strong correlation exists between PV power output and solar irradiance, compared with the atmospheric temperature and other meteorological variables. The correlation of other meteorological variables with the PV power output should be studied before designing the forecasting model. Therefore, the optimum number and most influential variables are recommended to establish a high-performance forecasting model to the PV power generation. Hence, the pattern and correlation of the meteorological variables and PV power output have been studied in numerous research.

2.2. Preprocessing of model input data

A significant number of forecasting models of the PV power generation have utilized the historical PV power output data and the meteorological data as their input. These historical data may contain different spikes and non-stationary components due to uncertain and changeable weather conditions. As a result, these spikes and non-stationary components in data will lead to a high forecast error due to the improper training of the models. In addition, the missing input data points in the historical data are due to the recording errors or other special events that will also play a role in increasing forecast error. The preprocessing of input data can significantly reduce the improper training problem and computational cost by learning the historical pattern properly. Therefore, the accuracy of the forecasting model can be considerably improved by the preprocessing of the input data.

A number of methods have been applied to preprocess the input data of forecasting models. Stationary, historical lag identification [28], trend-free time series [31], normalization [32], wavelet transform (WT) [33], and self-organizing map (SOM) [34] are useful methods to pre-process the input data. Among them, normalization [35–37] is widely used to preprocess the input data of the PV power generation forecasting model. In this technique, the input data are processed into a smaller range from a wider range of data. Hence, the data can be restricted within the range between 0 and 1 to minimize the regression error, improve precision, and maintain correlation among the dataset. The process formula is shown as follows:

$$I_{Normal} = \frac{I_{actual} - I_{min}}{I_{max} - I_{min}}$$

where I_{Normal} is the normalized data; I_{actual} is the original input data; and I_{min} and I_{max} are the minimum and maximum values of the input data, respectively.

If data are pre-processed, then post-processing is required ahead of the calculation or analysis on the performance of the forecasting model. In PV power forecasting, the most commonly used post-processing techniques are anti-normalization [32] and wavelet reconstruction [37]. If normalized data are used in the forecasting model, then the forecasted value should be anti-normalized to extract the actual forecasting PV power and analyze the performance of the model. Meanwhile, wavelet reconstruction is adopted to extract the actual forecasting PV power if the model input data are pre-processed by wavelet decomposition (WD).

3. Classification of PV power generation forecasting

Researchers have classified the forecasting of PV power generation in different categories based on different factors. However, no fixed criteria exist to classify the PV power forecasting. Most of the researchers classify the PV power forecasting depending on the forecasting horizon, historical data of solar irradiance and other meteorological data pattern, and methods used to forecast. An extensive review has been performed based on different classifications of PV power generation forecasting as presented in the following subsections.

3.1. Classification of PV power forecasting based on forecast horizon

The span of time into the future for which the PV power outputs are to be forecasted is called the forecast horizon. The purpose and accuracy of a PV power forecasting model depends on the forecast horizon. Lipperheide et al. [38] analyzed the performance of PV power forecasting over the different forecast horizons, such as 20, 40, 60, ... 180 s. The forecast error (RMSE) of the proposed forecasting model is in the range of 3.2–15.5% for forecast horizons from 20 s to 180 s. Lonij et al. [39] designed a PV power forecasting model where the errors have been changed with respect to forecast horizons ranging from 15 min to 90 min. The forecasting accuracy varies with the change of forecast horizon in the same model with the identical model parameters. Therefore, the forecast time horizon should be considered before designing the proper forecasting model. No well-defined criteria exist to classify the forecasting model based on the forecast horizon. Nevertheless, according to most of the researcher's reports, the forecasting of PV power generation can be divided into three categories based on time horizon, as shown in Fig. 6.

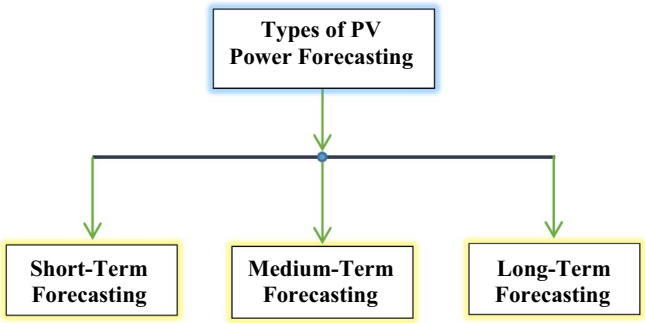


Fig. 6. Classification of PV power forecasting based on time horizon.

3.1.1. Short-term forecast

The forecasting of PV power generation done for one hour, several hours, one day, or up to seven days is known as short-term forecasting. The short-term forecasting of PV power generation ensures unit commitment, scheduling, and dispatching of electrical power. This type of forecasting model is useful in designing a PV-integrated energy management system. Short-term forecasting also enhances the security of grid operation.

3.1.2. Medium-term forecast

Medium-term PV power forecasting is done for more than one week to one month. This type of forecasting smooths the planning of the power system and maintenance schedule by predicting the availability of the electric power in the future.

3.1.3. Long-term forecast

Long-term forecasting of the PV power generation is done from one month to one year. This type of PV power forecasting is helpful for the planning of the electricity generation, transmission, and distribution organization aside from energy bidding and securing operation.

However, some of the researchers have divided the forecast horizon of PV power generation into four categories [40]. The fourth category is called “very short-term forecast horizon”. Very short-term PV power forecasting is considered for the few seconds, one minute, or several minutes (< 1 h) of forecast. This type of forecast has been done for power smoothing, real-time electricity dispatch, and optimal reserves.

3.2. Classification of PV power forecasting based on historical data

The forecasting methods of PV power generation can be categorized into four types based on the use of historical data of PV power output and related meteorological variables. These models are (a) persistence, (b) statistical, (c) machine-learning, and (d) hybrid method, as shown in Fig. 7 with subcategories.

In the persistence model, the forecasted PV power output is equal to the actual power output of the previous day at a similar time. In this method, only the historical PV power output data are required to forecast the PV power generation. Generally, this model is used as a benchmark model. In the statistical methods, the PV power generation is forecasted by the statistical analysis of the different input variables.

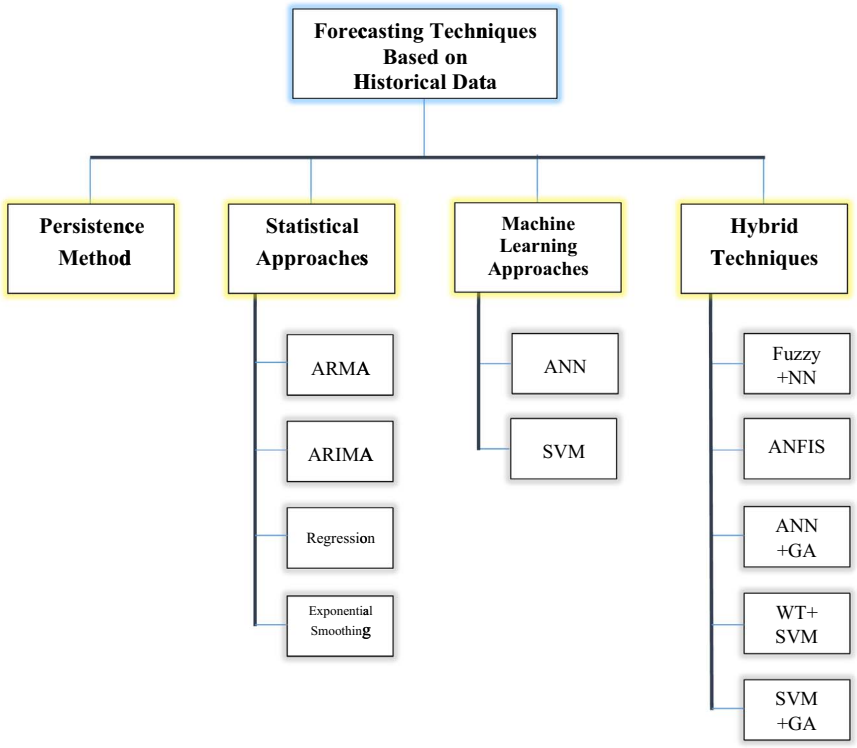


Fig. 7. Classification of PV power forecasting based on historical data.

Therefore, the past time-series data are used in these methods. Normally, these methods are adopted for the short-term forecasting of the PV power generation. The historical data at the recent time should be used in these methods for increasing model accuracy. The requirement of the input data series in this model is less compared to the machine-learning method. Conversely, in the machine-learning methods, a large data set is required to forecast the PV power generation accurately. The machine-learning model is an intelligent technique, and it can handle linear, non-linear, and non-stationary data patterns. The combination of two or more techniques is used to design a forecasting model known as the hybrid model. The hybrid model shows better results than a single model for different forecasting problems by combining the advantages of each individual technique.

3.3. Classification of PV power forecasting based on forecasting methods

Numerous researchers have studied the different methods for forecasting of PV power generation. They proposed and developed several techniques and models to forecast PV power generation. The most popular and significant forecasting methods are described as follows.

3.3.1. Persistence model

The persistence model is a considerably elementary forecasting model, which is commonly used to justify the performance of the other forecasting models. The forecasted value of any model is compared with the value of this model for accuracy evaluation. This model mainly acts as a benchmark for the other forecasting models. In this model, the forecasted PV power output is assumed to remain the same at the same time of the previous or following day. The forecasted PV power output for the next 24 h can be described as [41]:

$$P_f(t) = P_{pd}(t)$$

where P_f *is* the forecasted power, and P_{pd} is the output power of the previous day of the forecasted day at the same time t. For example, if the PV power output at 12:00 on February 21 needs to be forecasted, then this PV power output will be equal to the power at 12:00 on February 20.

This model is generally used for the short-term forecasting, especially the one-hour ahead forecasting of wind, solar, and PV powers. The forecasting error of this model depends on the stability of the weather condition. If the weather condition does not apparently change, then the PV power output of the previous day is a realistically a good indicator of the power output of the next day. However, the time range of forecasting increases, and the accuracy of this model decreases [42].

3.3.2. Autoregressive moving average (ARMA) model

The ARMA model consists of the combination of the two basic models: AR and MA models. ARMA is suitable for forecasting the PV power generation from the specified time-series data. The mathematical equation of the ARMA model is as follows [43]:

$$X(t) = \sum_{i=1}^{p} \alpha_i X(t-i) + \sum_{j=1}^{q} \beta_j e(t-j)$$

where $X(t)$ represents the forecasted PV power, which is the summation of the AR and MA functions. Hence, p and q indicates the order, α_i and β_i are the coefficients, of the AR and MA models, respectively. $e(t)$ is the white noise that produces random uncorrelated variables with zero mean and constant variance [44]. The main reason of the popularity of the ARMA model is its capability to extract the statistical properties and its adoption of the Box-Jenkins method [45]. An extension of this model is widely used for the different forecast models with an acceptable level of accuracy known as the AR integrated MA (ARIMA) [46]. In ARIMA model, an integrated part removes any non-stationarity from the data [47]. This model is to be used as a stand-

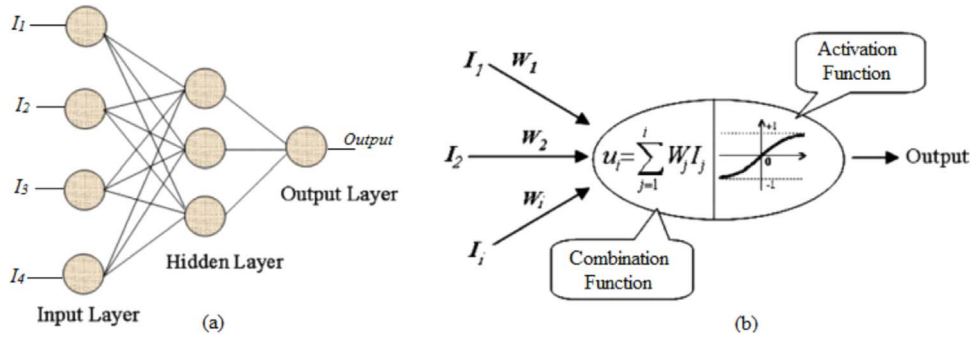


Fig. 8. (a) Schematic diagram of an ANN structure, which consists of an input, hidden, and output layers. (b) A mathematical model of an ANN cell [56].

alone or hybridized with other models to forecast the PV power generation. The main shortcoming of the ARMA model is that the used time series data must be stationary [48].

3.3.3. Regression method

The regression method is a statistical method used to establish a relationship between the explanatory and dependent variables. In this model, the dependent variable is forecasted by knowing the explanatory variables. In case of the PV power forecasting, the forecasted PV power is considered as a dependent variable, and the meteorological variables are considered as explanatory variables. Oudjana et al. forecasted the PV power generation, including two different regression models: the simple and multiple linear regressions [49]. The regression model using solar irradiance and temperature as input obtained better results compared to the case when either one was consider as input. A mathematical model and several explanatory variables are required to design a regression-based forecasting model, which is the weakness of this method.

3.3.4. Exponential smoothing method

The exponential smoothing method, which was first suggested by Brown [50], is known as the simple exponential smoothing method. This method was expanded by Holt, which is known as the Holt's method; Winter also modified this method as the Holt-Winter's method. In the exponential smoothing method, the assigned unequal set of weights of the past data instead of the equal weights is imposed to the past data. However, the weights of the past data decay exponentially from the most recent to the most distant data points. The simple exponential smoothing is a simple method, which is also known as the exponentially weighted MA (EWMA). The form of the EWMA model is assumed as follows:

$$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t = \hat{Y}_t + \alpha(Y_t - \hat{Y}_t)$$

where α is the smoothing constant, and it can assume any value between 0 and 1. This method requires an initial forecast that has to be either estimated or assumed. Hence, the initial forecast set $\hat{Y}_t$ is often $\hat{Y}_t = Y_1$. The EWMA forecasting model denotes that the forecasted value at time period $t+1$, is equal to the summation of the last forecasted value $(\hat{Y}_t)$and the adjustment for the error made in forecasting $\alpha(Y_t - \hat{Y}_t)$. Similar to the simple MA method, the EWMA method is appropriate for a stationary without trend time-series. The EWMA forecast would be lagging behind the trend if one exists.

3.3.5. Physical method

The physical forecasting method consists of the set of mathematical equations that describe the physical state and dynamic motion of the atmosphere [51]. Physical models are designed based on the PV power plant characteristics, such as location, different meteorological variables, and orientation historical data. These models can be considerably simple (when based only on the global solar irradiance) or complicated (if they include additional parameters) [52]. The accuracy of the physical forecasting model is higher when the weather conditions are stable [53]. However, the forecasting performance is largely affected due to the sharp changes in meteorological variables. Different studies [52,54] have reported that the physical forecasting method has been combined with the different artificial intelligence and statistic methods to design a hybrid forecasting model for better forecasting accuracy. However, these models are highly sensitive to weather forecasting and they should be designed specifically for a particular plant and location.

3.3.6. Artificial neural network (ANN)

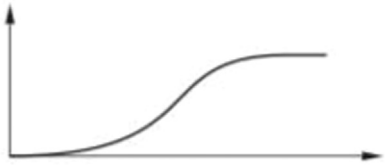
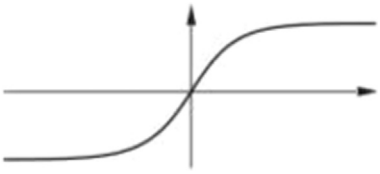
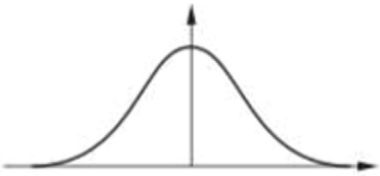
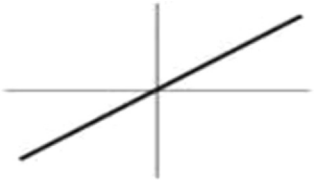
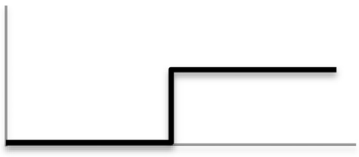
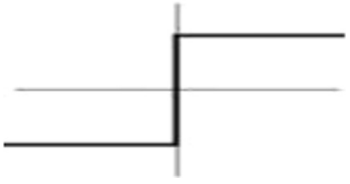
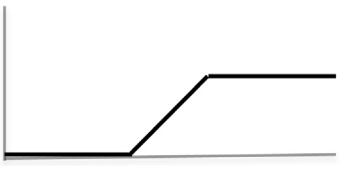
ANN is the most effective method and is popular among researchers since 1980. This method has been used in different prediction applications, including the forecasting of PV power generation with higher level of success. ANN is widely used in forecasting the PV power generation in most research because of non-linearity in meteorological data. ANN is more suitable compared with the statistical methods when a non-linear and complicated bonding exists between the data without any prior assumption.

The main components of an ANN are the input, hidden, and output layers, neurons, and connections. The input layer receives different input information. The hidden layer, which may consist of a single or several layers, analyzes the input information. The output layer receives the analyzed results and provides the output. The connection makes a link between the neurons of the different layers with update on its weight. Fig. 8(a) and(b) show the basic ANN architecture and a schematic model of the processing procedure in a neuron cell, respectively. The figure indicates that the neuron cell has two parts. The first part is the “combination function” that produces a value by summing up all the inputs. The second part is the “activation function”. The network output is generated by summing the weighted inputs using the activation function. Therefore, the activation function of the network acts as a squeezing function to transfer the input in the form of output. Different types of activation function are used in ANN. The most commonly used activation functions described in Table 1. The basic ANN mathematical formula expressed as [55]

$$U_N = b + \sum_{j=1}^{N} (W_j \times I_j)$$

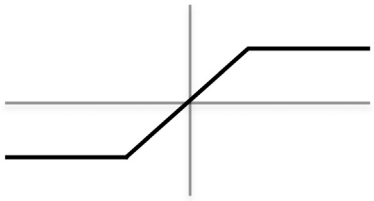
where, U_N,b, N, W_j, and I_j are the final network output, bias weight,

Table 1
Commonly used activation functions of ANN [28,57,58].

Name of the activation function	Formula	Graphical representation
Sigmoid	$f(u) = \frac{1}{1+e^{-u}}$	
Hyperbolic tangent sigmoid (tanh-sig)	$f(u) = tanh(u) = \frac{e^{u}-e^{-u}}{e^{u}+e^{-u}}$	
Gaussian radial basis	$f(u) = exp(-\|\|u-m\|\|^2/2\sigma^2)$	
Linear	$f(u) = u$	
Unipolar step function	$f(u) = \begin{cases} 0, \text{ if } u < 0 \\ 1, \text{ if } u > 0 \end{cases}$	
Bipolar step function	$f(u) = \begin{cases} -1, \text{ if } u < 0 \\ 0, \text{ if } u = 0 \\ 1, \text{ if } u > 0 \end{cases}$	
Unipolar linear function	$f(u) = \begin{cases} 0, \text{ if } u < 0 \\ y, \text{ if } u < 1 \\ 1, \text{ if } u > 1 \end{cases}$	
Bipolar linear function	$f(u) = \begin{cases} 1, \text{ if } u < -1 \\ y, \text{ if } \|u\| < 1 \\ 1, \text{ if } u > 1 \end{cases}$	

(continued on next page)

Table 1 *(continued)*

Name of the activation function	Formula	Graphical representation
		

number of input, connection weight, and network input, respectively. In most of the cases, the NN have multiple inputs and single or several outputs.

Training and testing are the two basic operations of an NN. At the training stage, the network is trained using the training input data set and learning algorithm. The learning algorithm of the NN attempts to map the input and output relationship by updating the synaptic weight values. The network generated output is compared with the desired output and then the error is calculated. Therefore, the weight and bias (acting as a compensator) values of NN are updated based on the error. This cycle continues until the desired output is achieved. Therefore, the network provides the final output based on the testing input data set and model weight value at the testing stage. However, the network output varies with the change in architecture, activation function, and NN inputs.

Among these activation functions, the most commonly used functions to forecast the PV power generations are the Gaussian radial basis, sigmoid, and hyperbolic tangent sigmoid functions because they are continuous, differentiable, and offer non-linearity to the network [59].

In most of the cases, the single-layer NNs have been used to solve the problem. However, a significant number of complex problems are caused by the unsolvable data pattern using a single-layer NN. Complex input and output relationships exist among different variables. In order to overcome the complex problem, NNs have been modified into several types which follow different architectures and different input-output mapping procedures. Among them, the most commonly used NNs are multilayer perceptron NN (MLPNN), multilayer feed-forward NN (MLFFNN), radial basis function NN (RBFNN), recurrent NN (RNN), general regression NN (GRNN), and adaptive neuro-fuzzy interface systems (ANFIS).

MLPNN is a supervised feed forward ANN [60] that consists of one or more hidden layers between the input and output layers of the network. The number of hidden layers can be adjusted in accordance with the complexity of a problem. A previous study revealed that more than two hidden layers is rarely required in performing complex mapping [61]. This NN has been adopted for different forecasting applications, such as PV power generation, electricity load and price forecasting [62]. MLFFNN is a relatively less complex NN architecture. In MLFNN, the information moves from the input layer to the output layer in a forward direction only. A feedback loop or cycle is not required in processing the information. This type of network has been adopted in several forecasting and pattern-recognition applications [63]. RBFNN is considered a two-layer NN, and the learning process can be divided into two different stages based on synaptic weights [64]. Mandal et al. reported that RBFNN involves less computation time for learning and achieves a good performance [65]. In PV power forecasting, RBFNN is selected for its structural simplicity and universal approximation property [66].

RNN performs well in learning different complex relationships and computational structures. Therefore, it is considered a good tool for time-series data forecasting [66]. Yona et al. reported that errors in forecasting results are significantly minimized by RNN compared with FFNN [67]. GRNN is a probabilistic-based network that performs regression rather than classification tasks [68]. With strong non-linear mapping capability, GRNN can effectively solve non-linear problems [69]. However, the growth in size of this method requires high computational power. The selection of network parameter σ is important in using GRNN [33]. Back-propagation (BP) algorithm is one of the most powerful supervised learning algorithms [70] however, this algorithm easily falls into the local minimum, has a slow convergence rate, and its training process is prone to oscillations [71]. The researchers introduced a modified BP network to address these limitations [72]. Among ANN-related methods, back propagation neural network (BPNN) has been widely used because of its excellent non-linear mapping function, which is suitable for solving complex

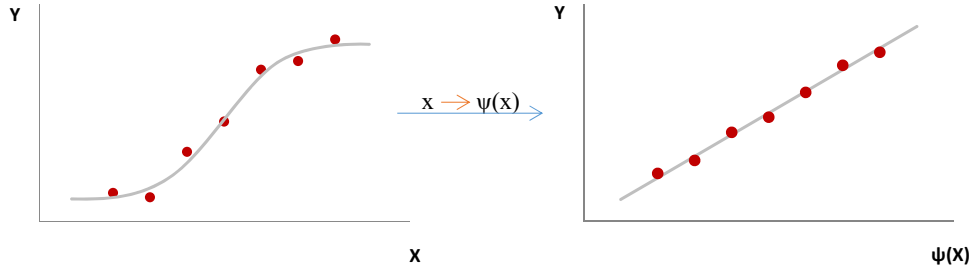


Fig. 9. Changing nonlinear regression into linear regression.

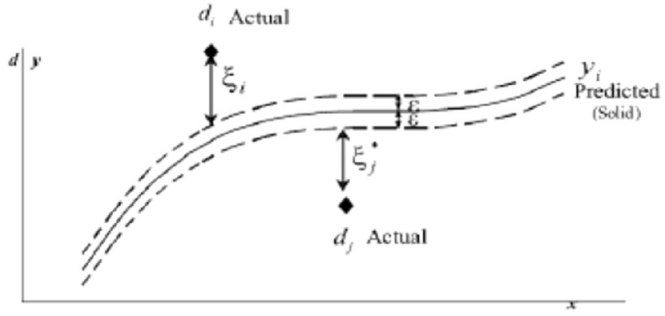


Fig. 10. Parameters used in support vector regression [78].

regression problems [73].

ANFIS is a type of adaptive multilayer feedforward network; it is applied to non-linear forecasting, in which past samples are used to forecast the sample in advance [74]. Tuning the membership functions is required to improve the performance of ANFIS [75]. Among fuzzy systems, ANFIS is most widely used because it is computationally less expensive, transparent, and produces results as robust as statistical models [76]. (See [58] for more information about different types of NNs).

However, the main drawbacks of ANN are requiring a large number of data during the training process and overfitting [52,77].

3.3.7. Support vector machine (SVM)

SVM [78] is a supervised machine-learning method based on structural risk minimization (SRM) principle. SRM minimizes an upper bound of the expected risk. Therefore, SVM can minimize the error of the training data. SVM was developed by Vapnik to solve the classification problem. However, SVM was recently extended to the domain of regression problems. The application of SVM in time series regression is known as support vector regression (SVR). The forecasting of PV power generation is a typical time series analysis problem; thus, SVR is a suitable method in this case [79].

The SVR algorithm is a nonlinear regression algorithm. The input time series data sample is mapped into a higher-dimensional feature space by nonlinear mapping, and then linear regression is performed in that space, as shown in Fig. 9.

Consider a set of training data, $\{(x_1, y_1), (x_2, y_2),\ldots\ldots,(x_l, y_l)\}$ where $x_l \in R^n$ is the input vector (meteorological variables data), and $y_l \in R^1$ is the corresponding output value (PV power output). The estimation function $f(x)$ is shown as [41,80]

$$y_f = f(x) = w \times \psi(x) + b$$

where $\psi(x)$ is the feature vector of inputs x; $w \in R^n$ is a weight vector, and $b \in R$ is the bias term, which are estimated by minimizing the regularized risk function [80,81]:

$$R(C) = C\frac{1}{N}\sum_{i=1}^{N} L_\varepsilon(y_i, f_i) + \frac{1}{2}\|W\|^2$$

where $L_\varepsilon(y_i, f_i) = \begin{cases} |y_i - f_i| - \varepsilon, & if\ |y_i - f_i| \geq \varepsilon \\ 0, & otherwise \end{cases}$ is the ε-insensitive loss function, C and ε are the user-determined prescribed parameters, y_i is the actual value at period i, and f_i is the forecasted value at period i. The loss will be equal to zero if the forecasted value is within the ε-tube (Fig. 10). The second term, $\frac{1}{2}\|W\|^2$ measures the flatness of the function. Two positive slack variables, ζ and ζ' represent the distance from actual values to the corresponding boundary values of the ε –tube.

Kernel function is a key factor of SVR. The performance of SVR highly depends on the selection of kernel function and its parameters [32]. There are many kernel functions that can be used in SVR; the most commonly used kernel are presented in Table 2.

Hence, x_i and x_j are inner vectors in the feature space $\phi(x_i)$ and $\phi(x_j)$ respectively, and γ, r and d are kernel parameters. The kernel parameters should be carefully chosen as they implicitly define the structure of the high dimensional feature space $\phi(x)$ and thus control the complexity of the final solution. The authors in [88–90] have discussed detail about the different kernel functions. RBF has only one parameter and a wide scope of applications [87]. Therefore, SVM with Gaussian RBF kernel function has been extensively used to forecast PV power generation.

To develop the SVR-based model for forecasting PV power generation, three parameters dominate the performance of the model [91]. These parameters are penalty (C), which determines the penalties for estimation errors, tube radius (ε), which determines the data inside the tube to be ignored in regression, and the kernel function's parameter. The suitable values of C, ε, and kernel function's parameter must be selected to develop the appropriate forecasting model. The performance of this model depends largely on the selection of the three parameters, which is a limitation of this method.

3.3.8. Hybrid model

The performance of a single model is not more accurate in forecasting the PV power generation in different cases. This problem can be attributed to the limitation of a stand-alone technique. The combination of two or more techniques, which is known as a hybrid model, is the best solution for those cases. These types of model have been used for several forecasting applications to achieve higher accuracy. One of the main motivations of these models is to explore the possible combinations of different topologies in improving the

Table 2
Most commonly used kernel function in SVR.

Function name	Function	Function explanation/applications	Reference
Liner	$k(x_i, x_j) = x_i^T x_j$.	Linear kernel function is applicable when the fluctuations of training data are small.	[82–84]
Polynomial	$k(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$.	This kernel represents the similarity of vectors (training samples) in a feature space. Thus, r, d, and γ are kernel parameters where d is the degree of polynomial.	[32,82,85–87]
Gaussian RBF	$k(x_i, x_j) = exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) = \exp(-\gamma\|x_i - x_j\|^2), \gamma > 0$.	This kernel is applicable to infinite dimensional feature space. The only kernel parameter, ($\gamma = 1/2\sigma^2$) is a constant that determines the width of the RBF kernel.	[32,82–85,87]
Sigmoid	$k(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$.	Thus, γ and r are two parameters. γ acts as a scaling parameter of the input data for γ > 0. r is a shifting parameter that controls the threshold of mapping.	[32,82,87]

forecast accuracy. The hybrid models can ensure the better forecasting performance by taking the advantage of each individual topology. Hybrid models have achieved better results than the stand-alone techniques for solving the PV power forecasting problem. Fuzzy inference model with RNN has been successfully applied in PV power forecasting [92]. In this case, fuzzy inference model has been utilized to smoothen the meteorological data, which is used for forecasting PV power generation. Hybrid fuzzy-GA forecasting model has also shown good results in PV power forecasting [93]. In several hybrid forecasting models, wavelet transform (WT) is used with conventional or artificial intelligence techniques in PV power forecasting. In such case, WT is used for de-noising the input data. The meteorological data after WT has been used as the input of ANN and SVM based forecasting models, which forecasted PV power generation with minimum error [94].

However, computational complexity is increased in a hybrid model due to the utilization of two or more techniques. The performance of a hybrid model depends on the performance of an individual single model. The accuracy of the hybrid model is affected by the selection of a single technique that may perform poorly, which is a limitation of the hybrid model.

4. Performance analysis of different forecasting methods

The persistence model, also known as the elementary model, has been extensively used as the benchmark for other simple and complex PV power forecasting models. This model can provide accurate forecasting of PV power generation for similar sunny weather conditions of neighboring forecasted days [95]. The primary weakness of this model is that the accuracy decreases significantly with the increase of cloud cover and when the duration is greater than 1 h [48]. Thus, this method is not suitable as a forecasting model; however, it provides a benchmark that can be widely used to standardize the forecasting performance of other models [96].

ARMA is an example of a linear method for the statistical modeling to correlate historical time-series data. This method can incorporate different types of time-series data and extract statistical properties from them. However, the time-series data need to be stationary. The ARMA method was improved with the introduction of the ARIMA method. ARIMA is an extension of ARMA with an added capability to handle non-stationary time-series data [97] and capture sharp transitions in irradiance [48] with greater accuracy. This capability to capture irregular data is desirable when the climate pattern being predicted is highly irregular. However, in comparison with ARMA, ARIMA is more computationally intensive due to the inclusion of a summation and integration function in the method.

ANN was developed as a shift away from linear algorithms to a non-linear solution. This method is remarkably popular due to its effectivity to be trained, which allows for greater accuracy. ANN has self-adaptive, fault-tolerance, robustness, and strong inference capabilities [34]. ANN-based modeling outperforms conventional mathematical models in terms of accuracy and adaptability [28,98]. The drawback with using ANN is that it significantly increases the complexity of the model due to the nature of using a multi-layered network architecture. In addition, NNs require a random initial data set that may reduce the reliability of the forecasted results.

Fuzzy logic (FL) provides a robust and advantageous method of modeling parameters and is based on inference rule. FL has been used for medium-term forecasting models, because it handles the uncertainty of real-world data [99]. In addition, fuzzy rules are linguistic-based and allow for a more statistical approach to decision making [100]. FL method has been combined with neural networking to produce an ANFIS method [101]. ANFIS combines FL with NN's self-adaptive training, which results in a 95% improvement in increase in accuracy from the standard FL method. The FL method, however, requires long execution time as opposed to SVM or several test nodes that limit its performance [102].

SVM is a prevailing and highly flexible non-linear model that has been used in wind power prediction and recently for forecasting of PV power generation. The strength of SVM is in its capability to learn without depending heavily on prior knowledge unlike ANN [103]. In addition, SVM does not have problems with local minima as with those of NN [104]. SVM allows for the simplification of the complex mathematical problems that may be related to PV forecasting [105]. However, SVM is highly sensitive to the parameters used to develop a forecasting model. These parameters are penalty factor (C), tube radius (ε), and kernel function parameter. The appropriate selection of these parameters is one of the difficulties of SVM.

Accuracy is a key point in the development of a forecasting model. The errors of the short-term forecasting of PV power generation should be less than 20% in RMSE [106]. The hybrid model has been extensively used to enhance the forecasting accuracy of PV power generation. The hybrid model has a number of examples to forecast PV power generation, as presented in Table 3 in the next section. In most of the cases, ANN and SVM-based models have adopted one or more different methods to obtain accurate forecasting of PV power generation. The authors developed a hybrid model based on weather classification method, in which different days are initially classified using SOM, LVQ, and similar day selection algorithm. The forecasting model was then designed to forecast PV power generation due to weather conditions. From different studies, the input data of the forecasting model were preprocessed by WT, selected by fuzzy system or k-means clustering, and smoothed by WT (that is unexpected spikes and sharp changes were removed), before being applied to the forecasting model. A hybrid model should be developed using superior attributes of two or more models for better forecasting performance. The poor performance of a single model in hybrid architecture may lead to higher forecasting error.

5. Optimization of forecasting model

A PV power forecasting model with high accuracy has to be developed to stabilize the grid operation and increase the penetration level of PV systems. The performance of forecasting models, especially machine-learning models, highly depends on several influential parameters. In ANN, the number of hidden layers, number of neurons, and epoch, are the selection parameters. In SVR, the penalty factor (C), tube radius (ε), and kernel function parameter are the hyperplane parameters. The performance of the forecasting model is strongly related to the selection of the aforementioned parameters. Therefore, the values of these parameters should be selected carefully to improve the accuracy of the model. However, these parameters are correlated with the suitable number of forecasting model inputs, data patterns, preprocessing of training/testing data of the network, and training techniques. Considering the above mentioned factors, the researchers selected the value of these parameters by trial-and-error method based

on their experience to enhance the accuracy of the forecasting model.

The selection of the appropriate number of neurons of the proposed NN-based model in Ref. [107] was an important task to forecast electric power demand. The author determined the appropriate number of hidden neurons by trial-and-error scheme to minimize the forecasting error. In Ref. [103], a feed-forward artificial NN was created with an adaptive BP method, where the number of hidden layers and the number of neurons in each layer are determined through trial-and-error method. Liu et al. [108] utilized an ANN approach to forecast the subsequent 24-h PV power outputs. In this study, a simple trial-and-error method was used to find the suitable number of neurons for balanced precision. The authors [109] selected the suitable parameter of SVR by using trial-and-error method to develop a SVR based electrical load forecasting model. Xu et al. [86] adopted a weighted-SVM to forecast the short-term PV power generation. The parameters C and ε were also set by trial-and-error method. Shi et al. [32] proposed an algorithm to forecast PV power output based upon weather classification and SVM. In this case, the dominated parameters of SVM were selected based on experience from a range of values. By contrast, the parameters (namely, fuzzy input variables and their membership functions) of fuzzy-based model were tuned by trial-and-error method on the training data set to achieve the best performance [34,110]. The selection of appropriate network parameters is crucial to this process. Therefore, the use of optimization algorithms is the best solution to identify the appropriate model parameters.

In many cases, optimization techniques have been used to determine the optimal value of different parameters. GA is one of the most popular evolution algorithms used to optimize weights and other features of forecasting methods. GA mimics the use of chromosomes to develop new DNA strings that combine elements from two different chromosome strands called crossover [111]. In [112], the weights and thresholds of BPNN were optimized using GA approach, which improved the forecasting accuracy of the proposed forecasting model. Pedro et al. [113] stated that the improvement of the ANN-based forecasting model can be achieved by GA optimization of the ANN parameters. In addition, several optimization algorithms, such as particle swarm optimization (PSO) [114], grid-search [115], fruit fly optimization algorithm (FOA) [116], firefly (FF) [117], ant colony optimization (ACO) [118], chaotic ant swarm optimization (CAS) [119], chaotic artificial bee colony algorithm [120], and immune algorithm (IA) [81], have been used in different applications for optimization. These optimization algorithms can be used to optimize the methods of PV power forecasting.

6. Performance evaluation matrices

The modern power grid system has been receiving a significant amount of electricity from PV power system, and this amount increases daily. As a result, the stability of grid systems has been highly dependent on the unit commitment of PV power systems. The grid system will be face difficulties in terms of providing adequate electricity supply if the PV system fails to supply committed electricity. Therefore, the accurate forecasting of PV power generation is important for ensuring grid stability and promoting PV installation. Thus, the accuracy measurement of the PV power forecasting model is a vital part of the forecasting process. Various evaluation matrices have been proposed and applied to measure the accuracy of PV power forecasting models. Standardized performance measures would help in forecasting model evaluation and benchmarking. MSE [86], RMSE [36], nRMSE [98], MAE [113], MAPE [108], MRE [121], and MBE [107], as shown below, have been commonly used in evaluating the accuracy of PV power forecasting models [36,86,98,107,108,113,121].

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(W_{forecasted} - W_{true})^2$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(W_{forecasted} - W_{true})^2}$$

$$nRMSE_{\%} = \left(\sqrt{\frac{1}{N}\sum_{i=1}^{N}(W_{forecasted} - W_{true})^2}\right)\times 100/W_{true\,(max)}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|W_{forecasted} - W_{true}|$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|W_{forecasting} - W_{true}|}{W_{true}}\times 100\%$$

$$MRE = \frac{1}{N}\sum_{i=1}^{N}\frac{W_{forecasting} - W_{true}}{W_{total}}\times 100\%$$

$$MBE = \frac{1}{N}\sum_{i=1}^{N}(W_{forecasted} - W_{true})$$

where, $W_{forecasting}$, W_{true}, $W_{true\,(max)}$, and W_{total} represent the forecasted PV power at each time point, the observed/measured PV power at each time point, the maximum observed PV power of this scale, and the PV installation capacity, respectively. N is the number of data sample for the time scale.

7. Analysis of recent works based on techniques and accuracy

Several studies have been conducted on PV power forecasting. A briefly description of the recent works on direct forecasting of PV power generation, including forecast horizon, accuracy, and methodology, are presented in Table 3.

Table 3
Recent works on PV power forecasting.

Authors and Ref.	Year	Forecast horizon	Forecast error	Forecasting model	Description of Main Works/contribution
Leva et al. [98]	2015	24 h ahead	Normalized RMSE (nRMSE) 12.5%–36.9%	ANN	ANN-based model was applied in this study to forecast the PV plant energy. The performance of the model during sunny, partially cloudy, and cloudy days was then analyzed. This study highlighted that the accuracy of the model is strictly related to the pre-processing of the historical data and the accuracy of the historical dataset used in the training step.
Liu et al. [108]	2015	24 h ahead	Mean absolute percentage error (MAPE) 7.65%	Back- propagation (BP) based ANN model	In this study, a novel PV power forecasting model was proposed considering aerosol index (AI) as an additional input parameter. AI was utilized along with seasonal weather classification and BP-based ANN model to forecast the next 24-h PV power generation. The results showed that the proposed model improved the prediction accuracy compared with the conventional ANN model due to the consideration of AI.
Chupong and Plangklang [35]	2011	1 d ahead	MAPE 16.83%	Elman NN	This study presented the PV power forecasting model using Elman NN. The calculated solar irradiance in a clear sky condition and weather forecast data were used as inputs of this model. The weather forecast data were collected from the PV system installed at the roof-top of an institutional building at Thanyaburi District in Thailand. The solar radiation measurement was avoided in this study due to high instrumentation cost.
Rana et al. [121]	2015	Every half-hour	Mean relative error (MRE) 16.92% –17.58%	NN, X-Means algorithm, and an iterative methodology	Three different approaches were proposed in this study based on the ensembles of NN to forecast the half-hourly PV power generation for the next day. Two of the proposed models were non-iterative, and one was iterative. The performance of the proposed models was evaluated using four Australian solar datasets for one year. The results revealed that the iterative approach is the most accurate approach compared with the others method. The authors also investigated the effect of using an ensemble NN instead of a single NN and showed that the use of an ensemble improved the forecasting accuracy.
Dolara et al. [122]	2015	24 h ahead	Normalized MAE (NMAE) < 1%, and weighted MAE (WMAE) < 2%	Physical model	Three physical models for forecasting the PV power generation by monocrystalline and polycrystalline PV panels were evaluated in this study. The comparison was performed using actual weather data measured by a meteorological station. The results showed that the accuracy of the model depends on (i) the data used for its calibration and (ii) the calculation of the cell temperature. Moreover, a training period was not required in this approach unlike ANN-based forecasting models. This approach is suitable for the initial period of a PV plant.
Tuyishimire et al. [123]	2013	1 d ahead	Not available	Kalman predictor	In this study, a forecasting model was developed based on multiple-rate Kalman predictor that provides real-time forecasting of PV power generation. This model used statistical data derived from historical data, which was correlated to daily and seasonal values. Two designs were compared to contrast the trade-off of steady-state variance versus transient-following capability.
Almeida et al. [107]	2015	1 d ahead	*cv*-Mean bias error (*cv* MBE) < 1.3%	Non−parametric model	A non-parametric forecasting model was proposed in this study to forecast the hourly AC power output of a PV plant 1 d ahead. This model used NWP model historical data and employed the PV system as a black box.
Kang et al. [124]	2011	24 h ahead	Approximate MAPE 11%	K-means clustering method	K-means clustering method was exploited at the first stage of this study. Five years historical data collected from a weather center were analyzed and classified according to rainfall probability patterns. The revised data were then applied to the PV power forecasting model, and the expected power generation was forecasted. The authors recommended the installation of additional equipment (digital camcorder) to observe the status of clouds. In this case, the error of weather forecasting data from the weather center could be modified to improve this algorithm.
Yang et al. [36]	2015	1–2 h ahead	RMSE- 81.5 for 1 h ahead and 136.3 for 2 h ahead	Spatial and temporal correlations	This study proposed a novel multiple time-scale data-driven forecasting model to improve the forecasting accuracy of short-term PV power generation. This model leveraged both spatial and temporal correlations among neighboring solar sites. The ARX-based ST model was used in this research to forecast PV power generation. The performance of the proposed model was evaluated using actual datasets from PV sites in California and Colorado. The results showed that this model outperformed the conventional persistence model.
Haque et al. [117]	2013	1 d ahead	MAPE 3.38%–11.83%, nRMSE 12.11% –13.13%	WT, fuzzy ARTMAP (FA), and firefly (FF)	The authors proposed a hybrid model to forecast PV power generation by classifying the season and forecasting the day based on weather conditions. WT was employed to

(continued on next page)

Table 3 *(continued)*

Authors and Ref.	Year	Forecast horizon	Forecast error	Forecasting model	Description of Main Works/contribution
					eliminate the unexpected spikes and sharp changes of input data. Furthermore, FA was used to forecast the PV power generation, and the model was optimized using an optimization technique based on FF algorithm.
Chen et al. [30]	2011	24 h ahead	MAPE 8.29%–10.80% (sunny day)	SOM and RBFNN	A new novel online forecasting model using RBFNN for forecasting of PV power generation was developed and applied in this study. The performance of the forecasting model varied with the type of day which the authors categorized into sunny, cloudy and rainy days using SOM. This model achieved good results in prediction of PV power generation during sunny and cloudy days. However, the forecasting result during rainy days was in acceptable range.
Ding et al. [71]	2011	24 h ahead	MAPE 10.06% (sunny day) and 18.89% (rainy day)	ANN and similar day selection algorithm	The standard BP algorithm has several limitations, such as slow convergence and easy to fall into the local minimum. The improved BP learning algorithm-based ANN model was adopted in this study to overcome the shortcomings of the standard BP learning algorithm. Similar day selection algorithm was adopted to improve the forecasting accuracy in different weather conditions using only historical power data and meteorological variable data. The proposed approach does not require complex modeling and complicated calculation.
Cococcioni et al. [125]	2011	24 h ahead	MAPE < 5.0%	Time series analysis and feed-forward NN (FFNN) with tapped delay lines	The authors implemented a 1 d ahead forecasting model to forecast PV energy production based on an ANN with tapped delay lines. A time-series analysis model NARX was utilized in this model. Given the focus of this research, a tool was proposed for the technician to configure the forecasting model correctly according to the particular installation characteristics. Experimental results of this research revealed that the performance of the proposed model is good.
De Giorgi et al. [126]	2014	1–24 h ahead	NMAE 6.50%–19.49%, nRMSE 10.91%–23.99%	Statistical methods based on multiple regression (MR) analysis and Elman ANN	In this study, a MR was used to analyze the sensitivity and impact of the different input vectors of the forecasting model. Elman ANN was employed to forecast PV power generation using three different input vectors as the training dataset. This model was applied to forecast the PV power generation of a grid connected PV plant installed in Italy. The results of this research revealed that the best performance of forecasting is found when all of the weather parameters, including PV power output data, are considered as the model input.
Tao and Chen [112]	2014	24 h ahead	Error 8.00%	GA-based NN	A distributed PV power forecasting method adopting the GA-based NN approach was proposed in this study. The weights and thresholds of back-propagation NN (BPNN) were optimized using the GA approach. This optimized model improved the forecasting accuracy of the BPNN model. The simulation results indicated that the proposed model outperformed BPNN-based forecasting model.
Chu et al. [127]	2015	5, 10, and 15 min ahead	Minimum mean absolute error (MAE) 21.02%	ANN with GA optimization	A GA optimized ANN-based real-time smart reforecasting model was developed in this study to enhance the intra-hour forecasting accuracy of PV power generation. The optimal set of input variables, number of hidden layers, and the number of neurons per layer was optimized by GA. The results indicated that the proposed model outperformed the three other models, namely, cloud tracking-based deterministic model, ARMA model, and a k-nearest neighbor (kNN) model.
Ramsami and Oree [128]	2015	24 h ahead	RMSE 2.74%	Stepwise regression, GRNN, FFNN, and MLR	In this study, stepwise regression model was used to select the input variables that were most strongly correlated with PV power generation. Then inputs were used for three single-stage forecasting models, namely, GRNN, FFNN, MLR model, and their corresponding hybrid models. The results revealed that the hybrid models performed slightly better than their corresponding single-stage models, and stepwise regression-FFNN hybrid model outperformed the other models.
Pedro and Coimbra [113]	2012	1-h and 2-h ahead	nRMSE 13.07 for 1-h and 18.71 for 2-h ahead	ANNs optimized by GA (GAs/ANN)	This study evaluated and compared five different forecasting techniques namely, persistent, ARIMA, kNN, ANN, and GAs/ANN) without the use of exogenous inputs, such as solar irradiance telemetry. The results showed that the accuracy of all the models strongly depends on the seasonal characteristics of solar variability. However, this study concluded that the ANN-based forecasting models outperformed the other models and that the optimization of ANN parameters by GA improved the forecasting accuracy.
Mandal et al. [65]	2012	1 h ahead	MAPE 2.38% (sunny day) and 4.08% (cloudy day)	WT and RBFNN	This study presented a PV power forecasting model using a combination of WT and RBFNN. This model incorporated the interactions of PV power output with solar

(continued on next page)

Table 3 *(continued)*

Authors and Ref.	Year	Forecast horizon	Forecast error	Forecasting model	Description of Main Works/contribution
					irradiance and temperature data. WT filtered the spike and chaotic changes in PV power and metrological time-series data. Meanwhile, RBFNN significantly captures the nonlinear PV power fluctuation. The results showed that the proposed model outperformed the single RBFNN model. However, the performance of the proposed model during rainy days was relatively unsatisfactory.
Yona et al. [92]	2013	24 h ahead	Maximum MAE 2.76 kW	Fuzzy theory (FT) and NN	FT and RNN-based models were proposed in this study to forecast PV power generation. This model considered the forecasted irradiance and weather report data. FT determined insolation forecast data for smoothing the training of the RNN. RNN was tested by PV power output data and weather reported data. The validity of the proposed method was verified by justifying the forecasting abilities on computer simulations.
Junior et al. [129]	2014	24 h ahead	RMSE 10.24%	SVR and principal component analysis (PCA)	In this study, three strategies based on past PV power generation data was evaluated to forecast 1 d ahead regional PV power generation using SVR, and identify the best strategy. These strategies were applied on six PV systems installed in different locations in Hokkaido, Japan. The results were showed that the regional forecast with preprocessing by PCA could significant improve the forecast error.
Yang et al. [34]	2014	1 d ahead	MRE 3.295%	SOM, learning vector quantization (LVQ), SVR, and fuzzy inference	This study presented a weather-based hybrid forecasting model to forecast the PV power generation. SOM and LVQ techniques were used to classify historical data based on weather conditions. SVR was used to train the forecasting model by appropriate dataset. In addition, fuzzy inference system was used to select a respective sub-model due to weather information for accurate forecasting. The proposed model outperformed the simple SVR and traditional ANN methods.
Shi et al. [32]	2012	1 d ahead	MRE 8.64%	Weather classification and SVM	The authors of this research proposed an algorithm to forecast PV power generation based on weather classification and SVM. The forecasted day was classified into four categories based on the weather conditions, namely, clear sky, foggy, cloudy, and rainy day. An SVM-based forecasting model with four sub-models was developed to forecast the PV power generation accurately. Each of these sub-models was trained based on the input vectors of similar category data. The sub-model was selected based on the weather condition of the forecasted day before forecasting the PV power generation.
Xu et al. [86]	2012	1 h ahead	Mean square error (MSE) 21.8	Similar day selection algorithm and weighted SVM (WSVM)	A model based on WSVM was developed in this study to forecast short-term PV power generation. In this case, five days with the most similarity to the day to be forecasted were selected as the training samples, and the weights of the samples for the WSVM were designed based on the similarities together with the time interval. The results showed that the proposed algorithm is valid and that the output of the proposed model is more efficient than that of the ANN.
Zhu et al. [37]	2015	1 d ahead	RMSE 7.193%–19.663%	Hybrid model (wavelet decomposition (WD) and ANN)	The authors proposed a hybrid model based on WD and ANN to forecast PV power generation in less calculation time. The historical PV power output was decomposed using WD to separate useful information from disturbances. ANN was used to develop the forecasting model with decomposed PV power output, including nonlinear relationships. Finally, the model output was reconstructed into the forecasted PV power.
Mori and Takahashi [130]	2012	24 h ahead	Maximum error 0.228 pu	Hybrid Method (GRBFN, DA, and evolutionary particle swarm optimization (EPSO))	In this study, deterministic annealing (DA) was used to determine the center and width of the RBF in generalized RBF network (GRBFN). Weight decay technique was employed to avoid overfitting for learning data of complicated non-linear time series. EPSO was used for optimizing weights between neurons in GRBFN, and GRBFN was used to forecast PV power generation. The simulation results presented that the proposed model significantly reduced the errors in comparison with conventional ANNs, such as MLP, RBFN, and GRBFN.

8. Conclusion

The grid integration of PV power has become a topic of research interest due to climate change and global warming in recent years. The accurate forecasting of PV power generation is essential because of the need to meet the increasing energy demand, mitigate climate change, and stabilize electric grid systems. Therefore, numerous studies have been conducted on this issue from different perspectives. In the present paper, a systematic and comprehensive literature review on direct PV power forecasting models and techniques was performed.

PV power generation and the necessity of PV power forecasting have been discussed at the beginning of this paper. In addition, the importance of the correlation of meteorological and PV power output data and the preprocessing of model input data have been considered. Findings show that the accuracy of a model increases when the inputs used in the model have high correlation factor with the output. Moreover, the accuracy of the forecasting result is enhanced when preprocessed input data are used. After discussing the input selection, a critical review on PV power forecasting were conducted based on the classification of the models from different perspectives. Findings show that the results of forecasting models depend on forecasting horizon, available data, and method employed for the forecasting. In most of the cases, the PV power forecasting model was developed based on short-term forecasting horizon. The accuracy of the PV power forecasting model varies by changing the forecast horizon, even with identical forecast model parameters.

In this paper, literature reviews of recent studies on direct forecasting of PV power generation, including the strengths and limitations of the models, have been presented as references for future researchers in this field. The performance of various recent forecasting methods were discussed and compared against different factors, such as accuracy, reliability, computational cost, and complexity. Based on recent studies, ANN and SVM-based forecasting models performed well under rapid and varying environmental conditions. In addition, the most of the studies adopted numerous techniques to develop the forecasting model for better accuracy. Moreover, a considerable number of studies classified the forecasted day into different categories based on the weather conditions using several techniques and then developed the forecasting model. However, the range of the observed error was remarkably high due to different weather conditions. The separate sub-model for each weather condition has to perform well to minimize errors. However the computational cost and complexity of the model should be considered in this case.

Moreover, the performance matrices used to evaluate the forecasting models have been introduced. Findings show that RMSE was used more frequently than others. Furthermore, the potential benefits of model optimization were discussed. Findings show that the optimized algorithm significantly increased the forecasting accuracy of the model. In this process, GA is one of the most viable optimization method in this field.

Acknowledgments

This work was supported by the Fundamental Research Grant Scheme (FRGS) (Project No: FP061-2016) and University of Malaya research grant (No. BK052-2016).

References

[1] Zervos A. Renewables 2016 Global Status Report; 2016.
[2] Zervos A, Lins C, Muth J. A 100% renewable energy vision for the European union-re-thinking 2050; 2010.
[3] Gueymard CA. The sun's total and spectral irradiance for solar energy applications and solar radiation models. Sol Energy 2004;76:423–53.
[4] Chen Z, Mo S, Hu P. Recent progress in thermodynamics of radiation—exergy of radiation, effective temperature of photon and entropy constant of photon. Sci China Ser E: Technol Sci 2008;51:1096–109.
[5] Goetzberger A, Hebling C, Schock H-W. Photovoltaic materials, history, status and outlook. Mater Sci Eng: R: Rep 2003;40:1–46.
[6] Shah ASBM, Yokoyama H, Kakimoto N. High-precision forecasting model of solar irradiance based on grid point value data analysis for an efficient photovoltaic system. IEEE Trans Sustain Energy 2015;6:474–81.
[7] Wengenmayr R, Bührke T. Renewable energy: sustainable energy concepts for the future. John Wiley & Sons; 2011.
[8] Jacobson MZ, Delucchi MA. Providing all global energy with wind, water, and solar power, Part I: technologies, energy resources, quantities and areas of infrastructure, and materials. Energy Policy 2011;39:1154–69.
[9] Byrne J, Kurdgelashvili L. The role of policy in PV industry growth: past, present and futureHandbook of photovoltaic science and engineering, 2nd ed. UK: John Wiley & Sons, Ltd.; 2011.
[10] Roselund C. The latest report by national renewable energy laboratory; 2016.
[11] Hoeven Mvd. Technology roadmap: solar photovoltaic energy. France: International Energy Agency; 2014. p. 20.
[12] Strzalka A, Alam N, Duminil E, Coors V, Eicker U. Large scale integration of photovoltaics in cities. Appl Energy 2012;93:413–21.
[13] Woyte A, Van Thong V, Belmans R, Nijs J. Voltage fluctuations on distribution level introduced by photovoltaic systems. IEEE Trans Energy Convers 2006;21:202–9.
[14] Tanaka K, Uchida K, Ogimi K, Goya T, Yona A, Senjyu T, et al. Optimal operation by controllable loads based on smart grid topology considering insolation forecasted error. IEEE Trans Smart Grid 2011;2:438–44.
[15] Zhang P, Takano H, Murata J. Daily solar radiation prediction based on wavelet analysis. SICE Annual Conference (SICE) Proceedings of: IEEE; 2011. p. 712–7.
[16] Yona A, Senjyu T, Funabshi T, Sekine H. Application of neural network to 24-hours-ahead generating power forecasting for PV system. IEEJ Trans Power Energy 2008;128:33–9.
[17] Capizzi G, Napoli C, Bonanno F. Innovative second-generation wavelets construction with recurrent neural networks for solar radiation forecasting. IEEE Trans Neural Netw Learn Syst 2012;23:1805–15.
[18] Cao S, Weng W, Chen J, Liu W, Yu G, Cao J Forecast of solar irradiance using chaos optimization neural networks. Power and Energy Engineering Conference, 2009 APPEEC 2009 Asia-Pacific: IEEE; 2009. p. 1–4.
[19] Mellit A, Pavan AM. A 24-h forecast of solar irradiance using artificial neural network: application for performance prediction of a grid-connected PV plant at Trieste, Italy. Sol Energy 2010;84:807–21.
[20] Wang F, Mi Z, Su S, Zhao H. Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters. Energies 2012;5:1355–70.
[21] Hocaoğlu FO, Gerek ÖN, Kurban M. Hourly solar radiation forecasting using optimal coefficient 2-D linear filters and feed-forward neural networks. Sol Energy 2008;82:714–26.
[22] Dalton G, Lockington D, Baldock T. Feasibility analysis of renewable energy supply options for a grid-connected large hotel. Renew Energy 2009;34:955–64.
[23] Kudo M, Takeuchi A, Nozaki Y, Endo H, Sumita J. Forecasting electric power generation in a photovoltaic power system for an energy network. Electr Eng Jpn 2009;167:16–23.
[24] Ren Y, Suganthan P, Srikanth N. Ensemble methods for wind and solar power forecasting—a state-of-the-art review. Renew Sustain Energy Rev 2015;50:82–91.
[25] Gandoman FH, Raeisi F, Ahmadi A. A literature review on estimating of PV-array hourly power under cloudy weather conditions. Renew Sustain Energy Rev 2016;63:579–92.
[26] Antonanzas J, Osorio N, Escobar R, Urraca R, Martinez-de-Pison F, Antonanzas-Torres F. Review of photovoltaic power forecasting. Sol Energy 2016;136:78–111.
[27] Wan C, Zhao J, Song Y, Xu Z, Lin J, Hu Z. Photovoltaic and solar power forecasting for smart grid energy management. CSEE J Power Energy Syst 2015;1:38–46.
[28] Raza MQ, Nadarajah M, Ekanayake C. On recent advances in PV output power forecast. Sol Energy 2016;136:125–44.
[29] Huang Y, Lu J, Liu C, Xu X, Wang W, Zhou X. Comparative study of power forecasting methods for PV stations. In: Proceedings of international conference on power system technology (POWERCON). IEEE; 2010. p. 1–6.
[30] Chen C, Duan S, Cai T, Liu B. Online 24-h solar power forecasting based on weather type classification using artificial neural network. Sol Energy 2011;85:2856–70.
[31] Azadeh A, Ghaderi S, Sohrabkhani S. Forecasting electrical consumption by integration of neural network, time series and ANOVA. Appl Math Comput 2007;186:1753–61.
[32] Shi J, Lee W-J, Liu Y, Yang Y, Wang P. Forecasting power output of photovoltaic systems based on weather classification and support vector machines. IEEE Trans Ind Appl 2012;48:1064–9.
[33] AlHakeem D, Mandal P, Haque AU, Yona A, Senjyu T, Tseng T-L. A new strategy to quantify uncertainties of wavelet-GRNN-PSO based solar PV power forecasts using bootstrap confidence intervals. Power & Energy Society General Meeting IEEE: IEEE; 2015. p. 1–5.
[34] Yang H-T, Huang C-M, Huang Y-C, Pai Y-S. A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output. IEEE Trans Sustain Energy 2014;5:917–26.
[35] Chupong C, Plangklang B. Forecasting power output of PV grid connected system in Thailand without using solar radiation measurement. Energy Procedia 2011;9:230–7.
[36] Yang C, Thatte AA, Xie L. Multitime-scale data-driven spatio-temporal forecast of photovoltaic generation. IEEE Trans Sustain Energy 2015;6:104–12.
[37] Zhu H, Li X, Sun Q, Nie L, Yao J, Zhao G. A power prediction method for photovoltaic power plant based on wavelet decomposition and artificial neural networks. Energies 2015;9:11.

[38] Lipperheide M, Bosch J, Kleissl J. Embedded nowcasting method using cloud speed persistence for a photovoltaic power plant. Sol Energy 2015;112:232–8.
[39] Lonij VP, Brooks AE, Cronin AD, Leuthold M, Koch K. Intra-hour forecasts of solar power production using measurements from a network of irradiance sensors. Sol Energy 2013;97:58–66.
[40] Amral N, Ozveren C, King D Short term load forecasting using multiple linear regression. In: Proceeding of the 42nd international universities power engineering conference, 2007. IEEE; 2007. p. 1192–8.
[41] Yang X, Ren J, Yue H. Photovoltaic power forecasting with a rough set combination method. In: Proceedings of the UKACC 11th international conference on control (CONTROL). IEEE; 2016. p. 1–6.
[42] Perez R, Kivalov S, Schlemmer J, Hemker K, Renné D, Hoff TE. Validation of short and medium term operational solar radiation forecasts in the US. Sol Energy 2010;84:2161–72.
[43] Huang R, Huang T, Gadh R, Li N. Solar generation prediction using the ARMA model in a laboratory-level micro-grid. in: IEEE proceedings of the third international conference on smart grid communications (SmartGridComm) IEEE; 2012. p. 528–33.
[44] Rajagopalan S, Santoso S. Wind power forecasting and error analysis using the autoregressive moving average modeling. Power & Energy Society General Meeting 2009PES'09 IEEE: IEEE; 2009. p. 1–6.
[45] Boland J. Time series modelling of solar radiationModeling solar radiation at the earth's surface. Springer; 2008. p. 283–312.
[46] Box GE, Jenkins GM, Reinsel GC, Ljung GM. Time series analysis: forecasting and control. John Wiley & Sons; 2015.
[47] Wan Ahmad WKA, Ahmad S, Ishak A, Hashim I, Ismail ES, Nazar R. Arima model and exponential smoothing method: a comparison. AIP Conference Proceedings: AIP; 2013. p. 1312–21.
[48] Diagne M, David M, Lauret P, Boland J, Schmutz N. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renew Sustain Energy Rev 2013;27:65–76.
[49] Oudjana SH, Hellal A, Mahamed IH. Short term photovoltaic power generation forecasting using neural network. In: Proceedings of the 11th international conference on environment and electrical engineering (EEEIC) 2012. IEEE; 2012. p. 706–11.
[50] Brown RG. Exponential smoothing for predicting demand. Operations research: inst operations research management sciences 901 elkridge landing rd, Ste 400, Linthicum Hts, MD 21090-2909; 1957. p. 145–145.
[51] Monteiro C, Fernandez-Jimenez LA, Ramirez-Rosado IJ, Muñoz-Jimenez A, Lara-Santillan PM. Short-term forecasting models for photovoltaic plants: analytical versus soft-computing techniques. Math Probl Eng 2013;2013.
[52] Dolara A, Grimaccia F, Leva S, Mussetta M, Ogliari E. A physical hybrid artificial neural network for short term forecasting of PV plant power output. Energies 2015;8:1138–53.
[53] Soman SS, Zareipour H, Malik O, Mandal P. A review of wind power and wind speed forecasting methods with different time horizons. North American Power Symposium (NAPS), 2010 IEEE. 2010. p. 1–8.
[54] Gandelli A, Grimaccia F, Leva S, Mussetta M, Ogliari E Hybrid model analysis and validation for PV energy production forecasting. Neural Networks (IJCNN), 2014 International Joint Conference on: IEEE; 2014. p. 1957–1962.
[55] Hossain MS, Chao OZ, Ismail Z, Noroozi S, Yee KS. Artificial neural networks for vibration based inverse parametric identifications: a review. Appl Soft Comput 2016.
[56] Aminzadeh F, De Groot P. Neural networks and other soft computing techniques with applications in the oil industry. Eage Publications; 2006.
[57] Hu YH, Hwang J-N. Handbook of neural network signal processing. CRC Press; 2001.
[58] Raza MQ, Khosravi A. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. Renew Sustain Energy Rev 2015;50:1352–72.
[59] Mellit A, Kalogirou SA. Artificial intelligence techniques for photovoltaic applications: a review. Progress Energy Combust Sci 2008;34:574–632.
[60] Isa IS, Omar S, Saad Z, Noor NM, Osman MK. Weather forecasting using photovoltaic system and neural network. Computational Intelligence, Communication Systems and Networks (CICSyN) s International Conference on: IEEE. 2010. p. 96–100.
[61] Zhang Y, Chen G, Malik O, Hope G. An artificial neural network based adaptive power system stabilizer. IEEE Trans Energy Convers 1993;8:71–7.
[62] Tasre MB, Bedekar PP, Ghate VN. Daily peak load forecasting using ANN. In: Proceedings of Nirma University International Conference on Engineering (NUiCONE) 2011. IEEE; 2011. p. 1–6.
[63] Malki HA, Karayiannis NB, Balasubramanian M. Short-term electric power load forecasting using feedforward neural networks. Expert Syst 2004;21:157–67.
[64] Gupta M, Jin L, Homma N. Static and dynamic neural networks: from fundamentals to advanced theory. John Wiley & Sons; 2004.
[65] Mandal P, Madhira STS, Meng J, Pineda RL. Forecasting power output of solar photovoltaic system using wavelet transform and artificial intelligence techniques. Procedia Comput Sci 2012;12:332–7.
[66] Yona A, Senjyu T, Saber AY, Funabashi T, Sekine H, Kim C-H Application of neural network to one-day-ahead 24h generating power forecasting for photovoltaic system. In: Proceedings of international conference on intelligent systems applications to power systems, ISAP 2007. IEEE; 2007. p. 1–6.
[67] Yona A, Senjyu T, Funabashi T. Application of recurrent neural network to short-term-ahead generating power forecasting for photovoltaic system. 2007 IEEE Power Engineering Society General Meeting; 2007. p. 1–6.
[68] Ibrić S, Djuriš J, Parojčić J, Djurić Z. Artificial neural networks in evaluation and optimization of modified release solid dosage forms. Pharmaceutics 2012;4:531–50.
[69] Li H-Z, Guo S, Li C-J, Sun J-Q. A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm. Knowl-Based Syst 2013;37:378–87.
[70] Sözen A, Arcaklioğlu E, Özalp M, Kanit EG. Use of artificial neural networks for mapping of solar potential in Turkey. Appl Energy 2004;77:273–86.
[71] Ding M, Wang L, Bi R. An ANN-based approach for forecasting the power output of photovoltaic system. Procedia Environ Sci 2011;11:1308–15.
[72] Bizzarri F, Bongiorno M, Brambilla A, Gruosso G, Gajani GS. Model of photovoltaic power plants for performance analysis and production forecast. IEEE Trans Sustain Energy 2013;4:278–85.
[73] Smolensky P, Mozer MC, Rumelhart DE. Mathematical perspectives on neural networks. Psychology Press; 2013.
[74] Yun Z, Quan Z, Caixin S, Shaolan L, Yuming L, Yang S. RBF neural network and ANFIS-based short-term load forecasting approach in real-time price environment. IEEE Trans Power Syst 2008;23:853–8.
[75] Catalao J, Pousinho H, Mendes V. Hybrid wavelet-PSO-ANFIS approach for short-term wind power forecasting in Portugal. IEEE Trans Sustain Energy 2011;2:50–9.
[76] Awadallah MA, Bayoumi EH, Soliman HM. Adaptive deadbeat controllers for brushless DC drives using PSO and ANFIS techniques. J Electr Eng 2009;60:3–11.
[77] De Giorgi M, Malvoni M, Congedo P. Comparison of strategies for multi-step ahead photovoltaic power forecasting models based on hybrid group method of data handling networks and least square support vector machine. Energy 2016;107:360–73.
[78] Müller K-R, Smola AJ, Rätsch G, Schölkopf B, Kohlmorgen J, Vapnik V. Predicting time series with support vector machines. International Conference on Artificial Neural Networks: Springer. 1997. p. 999–1004.
[79] Hu J, Gao P, Yao Y, Xie X. Traffic flow forecasting with particle swarm optimization and support vector regression. In: Proceedings of the IEEE 17th international conference on intelligent transportation systems (ITSC); 2014. p. 2267–2268.
[80] Zhang F, Deb C, Lee SE, Yang J, Shah KW. Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique. Energy Build 2016;126:94–103.
[81] Hong W-C. Electric load forecasting by support vector model. Appl Math Model 2009;33:2444–54.
[82] Tang Y, Zhou J. The performance of PSO-SVM in inflation forecasting. In: Proceedings of the 12th international conference on service systems and service management (ICSSSM) IEEE; 2015. p. 1–4.
[83] Mao M, Gong W, Chang L. Short-term photovoltaic output forecasting model for economic dispatch of power system incorporating large-scale photovoltaic plant. Energy Convers Congr Expo (ECCE) 2013;2013:4540–5, [IEEE: IEEE].
[84] Lin K-P, Pai P-F. Solar power output forecasting using evolutionary seasonal decomposition least-square support vector regression. J Clean Prod 2016;134:456–62.
[85] Yang X, Jiang F, Liu H. Short-term solar radiation prediction based on SVM with similar data; 2013.
[86] Xu R, Chen H, Sun X. Short-term photovoltaic power forecasting with weighted support vector machine. In: Proceedings of IEEE international conference on automation and logistics (ICAL); 2012. p. 248–53.
[87] Gao P, Hu J, Zhou H, Zhang Y. Travel time prediction with immune genetic algorithm and support vector regression. In: Proceedings of 12th world congress on intelligent control and automation (WCICA), 2016. IEEE; 2016. p. 987–92.
[88] Samsudin R, Saad P, Shabri A. A hybrid GMDH and least squares support vector machines in time series forecasting. Neural Netw World 2011;21:251.
[89] Huang H-Y, Lin C-J. Linear and kernel classification: when to use which? In: Proceedings of the 2016 siam international conference on data mining: SIAM; 2016. p. 216–24.
[90] Lin H-T, Lin C-J. A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. Neural Comput 2003:1–32, [submitted to].
[91] Cherkassky V, Ma Y. Practical selection of SVM parameters and noise estimation for SVM regression. Neural Netw 2004;17:113–26.
[92] Yona A, Senjyu T, Funabashi T, Kim C-H. Determination method of insolation prediction with fuzzy and applying neural network for long-term ahead PV power output correction. IEEE Trans Sustain Energy 2013;4:527–33.
[93] Chen S-M, Chang Y-C, Pan J-S. Fuzzy rules interpolation for sparse fuzzy rule-based systems based on interval type-2 Gaussian fuzzy sets and genetic algorithms. IEEE Trans Fuzzy Syst 2013;21:412–25.
[94] Colak T, Qahwaji R. Automatic sunspot classification for real-time forecasting of solar activities. In: Proceedings of the 3rd international conference on recent advances in space technologies 2007 RAST'07. IEEE; 2007. p. 733–8.
[95] Xiyun Y, Song C. A combination method in photovoltaic power forecasting based on the correlation coefficient. In: Proceedings of the 5th IEEE international conference on software engineering and service science (ICSESS) 2014. IEEE; 2014. p. 706–9.
[96] Yadav HK, Pal Y, Tripathi M. Photovoltaic power forecasting methods in smart power grid. India Conference (INDICON) Annual IEEE: IEEE; 2015. p. 1–6.
[97] Reikard G. Predicting solar radiation at high resolutions: a comparison of time series forecasts. Sol Energy 2009;83:342–9.
[98] Leva S, Dolara A, Grimaccia F, Mussetta M, Ogliari E. Analysis and validation of 24 h ahead neural network forecasting of photovoltaic output power. Math Comput Simul 2017;131:88–100.
[99] Ghofrani M, Ghayekhloo M, Azimi R. A novel soft computing framework for solar radiation forecasting. Appl Soft Comput 2016;48:207–16.

[100] Gupta A, Chauhan YK, Pachauri RK. A comparative investigation of maximum power point tracking methods for solar PV system. Sol Energy 2016;136:236–53.
[101] Law EW, Prasad AA, Kay M, Taylor RA. Direct normal irradiance forecasting and its application to concentrated solar thermal output forecasting–a review. Sol Energy 2014;108:287–307.
[102] Chen P, Yuan L, He Y, Luo S. An improved SVM classifier based on double chains quantum genetic algorithm and its application in analogue circuit diagnosis. Neurocomputing 2016;211:202–11.
[103] Li Y, He Y, Su Y, Shu L. Forecasting the daily power output of a grid-connected photovoltaic system based on multivariate adaptive regression splines. Appl Energy 2016;180:392–401.
[104] Lauret P, Voyant C, Soubdhan T, David M, Poggi P. A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. Sol Energy 2015;112:446–57.
[105] Mojumder JC, Ong HC, Chong WT, Shamshirband S. Application of support vector machine for prediction of electrical and thermal performance in PV/T system. Energy Build 2016;111:267–77.
[106] Daye T. Managing intermittency: standards and recommended practices in solar power forecasting. Austin: ERCOT Emerging Technol Working Group; 2011.
[107] Almeida MP, Perpiñán O, Narvarte L. PV power forecast using a nonparametric PV model. Sol Energy 2015;115:354–68.
[108] Liu J, Fang W, Zhang X, Yang C. An improved photovoltaic power forecasting model with the assistance of aerosol index data. IEEE Trans Sustain Energy 2015;6:434–42.
[109] Nagi J, Yap K, Tiong S, Ahmed S. Electrical power load forecasting using hybrid self-organizing maps and support vector machines. Training 2008;99:31.
[110] Huang C-MT, Huang Y-C, Huang K-Y. A hybrid method for one-day ahead hourly forecasting of PV power output. In: Proceedings of the 9th conference on industrial electronics and applications (ICIEA) 2014. IEEE; 2014. p. 526–31.
[111] Qiongbing Z, Lixin D. A new crossover mechanism for genetic algorithms with variable-length chromosomes for path optimization problems. Expert Syst Appl 2016;60:183–9.
[112] Tao Y, Chen Y Distributed PV. power forecasting using genetic algorithm based neural network approach. In: Proceedings of international conference on advanced mechatronic systems (ICAMechS) 2014. IEEE; 2014. p. 557–60.
[113] Pedro HT, Coimbra CF. Assessment of forecasting techniques for solar power production with no exogenous inputs. Sol Energy 2012;86:2017–28.
[114] Bashir Z, El-Hawary M. Applying wavelets to short-term load forecasting using PSO-based neural networks. IEEE Trans Power Syst 2009;24:20–7.
[115] Bao Y, Liu Z. A fast grid search method in support vector regression forecasting time series. In: Proceedings of international conference on intelligent data engineering and automated learning: Springer; 2006. p. 504–11.
[116] Li H, Guo S, Zhao H, Su C, Wang B. Annual electric load forecasting by a least squares support vector machine with a fruit fly optimization algorithm. Energies 2012;5:4430–45.
[117] Haque AU, Nehrir MH, Mandal P Solar PV. power generation forecast using a hybrid intelligent approach. Power and Energy Society General Meeting (PES) IEEE: IEEE; 2013. p. 1–5.
[118] Niu D, Wang Y, Wu DD. Power load forecasting using support vector machine and ant colony optimization. Expert Syst Appl 2010;37:2531–9.
[119] Hong W-C. Application of chaotic ant swarm optimization in electric load forecasting. Energy Policy 2010;38:5830–9.
[120] Hong W-C. Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm. Energy 2011;36:5568–78.
[121] Rana M, Koprinska I, Agelidis VG. Forecasting solar power generated by grid connected PV systems using ensembles of neural networks. In: Proceedings of international joint conference on neural networks (IJCNN) 2015. IEEE; 2015. p. 1–8.
[122] Dolara A, Leva S, Manzolini G. Comparison of different physical models for PV power output prediction. Sol Energy 2015;119:83–99.
[123] Tuyishimire B, McCann R, Bute J. Evaluation of a Kalman predictor approach in forecasting PV solar power generation. In: Proceedings of the 4th IEEE international symposium on power electronics for distributed generation systems (PEDG); 2013. p. 1–6.
[124] Kang M-C, Sohn J-M, Park J-y, Lee S-K, Yoon Y-T. Development of algorithm for day ahead PV generation forecasting using data mining method. In: IEEE Proceedings of the 54th international midwest symposium on circuits and systems (MWSCAS) 2011; 2011. p. 1–4.
[125] Cococcioni M, D'Andrea E, Lazzerini B. 24-hour-ahead forecasting of energy production in solar PV systems. In: Proceedings of the 11th international conference on intelligent systems design and applications (ISDA) 2011. IEEE; 2011. p. 1276–81.
[126] De Giorgi M, Congedo P, Malvoni M. Photovoltaic power forecasting using statistical methods: impact of weather data. IET Sci Meas Technol 2014;8:90–7.
[127] Chu Y, Urquhart B, Gohari SM, Pedro HT, Kleissl J, Coimbra CF. Short-term reforecasting of power output from a 48 MWe solar PV plant. Sol Energy 2015;112:68–77.
[128] Ramsami P, Oree V. A hybrid method for forecasting the energy output of photovoltaic systems. Energy Convers Manag 2015;95:406–13.
[129] Junior JGdSF, Oozeki T, Ohtake H, Shimose K-i, Takashima T, Ogimoto K. Energy Procedia 2014;57:1337–45.
[130] Mori H, Takahashi M. Development of GRBFN with global structure for PV generation output forecasting. Power and energy society general meeting 2012 IEEE: IEEE; 2012. p. 1–7.